ARIC Manuscript Proposal # 3307

PC Reviewed: 12/11/18 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: Working Title: Putative Markers for the Detection of Physical Frailty: A Population Predictive Model

b. Abbreviated Title (Length 26 characters):
Predicting Frailty: A Population Predictive Model

2. Writing Group:
Writing group members: Lana Sargent, Ph.D., RN, CRNP, Andrew Singleton, Ph.D., Mike Nalls, Ph.D., Luigi Ferrucci, M.D., Ph.D., Priya Palta PhD, Anna Kucharska-Newton PhD, MPH, Rajesh Talluri Ph.D., Jim Pankow, Ph.D., Hunter Young, M.D., M.H.S., Weihong Tang, Ph.D., Danni Li, Ph.D., Pamela Lutsey, Ph.D., Alvaro Alonso, M.D., Ph.D., Michael Griswold Ph.D., B. Gwen Windham, M.D., MHS

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. _LS____ **[please confirm with your initials electronically or in writing]**

First author: Lana Sargent, Ph.D., RN, CRNP
Address: Molecular Genetics Section and Laboratory of Neurogenetics, NIA,NIH
Building 35, Room 1A1014, 35 Convent Drive, Bethesda, MD 20892
Phone: 781-548-9128 Fax: 301-451-5466
E-mail: lana.sargent@nih.gov

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).
Name: Beverly Gwen Windham, M.D., MHS
Address: 2500 N. State St. Jackson, MS 39216
Phone: 601-984-5645 Fax: 601-984-5783
E-mail: gwindham@umc.edu

3. Timeline:

Study Timeline	Year	One					
	Month	2	4	6	8	10	12
Cleaning Data: Identify phenotype and biomarkers and organize data frame							
Data analysis: Train the predictive classification models, tune xgboost to maximize testing							
Model evaluation and dissemination/publication							

4. Rationale:

We seek to complete and develop a population predictive model in ARIC using data from a frailty predictive model that was developed in the longitudinal study in aging *Invecchaiare in Chianti* (Aging in Chianti, "InCHIANTI Study") **See background and significance**

Background

There have been significant changes to the age demographics in the U.S. with Americans 65 and older reaching more than 20% of the total population by 2030.[1] Age demographic changes have led to increased hospital utilization by older adults. The National Hospital Discharge Survey demonstrated hospital utilization by older adults increased from 20% to 37% with older adults accounting for 43% of the days of care.[2] Older adults present with a unique biophysical profile making them more vulnerable to acute and chronic health exacerbations and postoperative complications resulting in subsequent rapid decline in physical and cognitive functioning.

Frailty is a phenotype of physiological reserve and can be used to measure resistance to stressors.[3,4] The frailty phenotype has increasingly been recognized as a validated measure of decreased physiological reserve and predictor of poor outcomes in older adults in medical and surgical specialties.[4] Frailty is one of the greatest challenges for health care professionals with aging populations. It is associated with adverse health outcomes, dependency, and increased rates of hospital admission and mortality.[3,5,6] Additionally, frail patients have significantly higher 1-year postoperative mortality rates.[4] Perioperative pre-frail and frail older adult patients have a 2.06 [95% CI, 1.18-3.60] and 2.54 times higher odds [1.12-5.77] of postoperative adverse events; increased length of stay, and higher likelihood of discharge to a skilled or assisted-living facility.[4,7] This increased risk spurred a joint statement from the American College of Surgeons and the American Geriatrics Society in 2012 to recommend a frailty assessment as a part of the preoperative evaluation for all older adults.[4] Subsequently, the Society for Perioperative Assessment and Quality Improvement (SPAQI) outlined practical steps for clinicians to asses frailty in older adults who require elective intermediate or high risk surgery.[7]

Significance

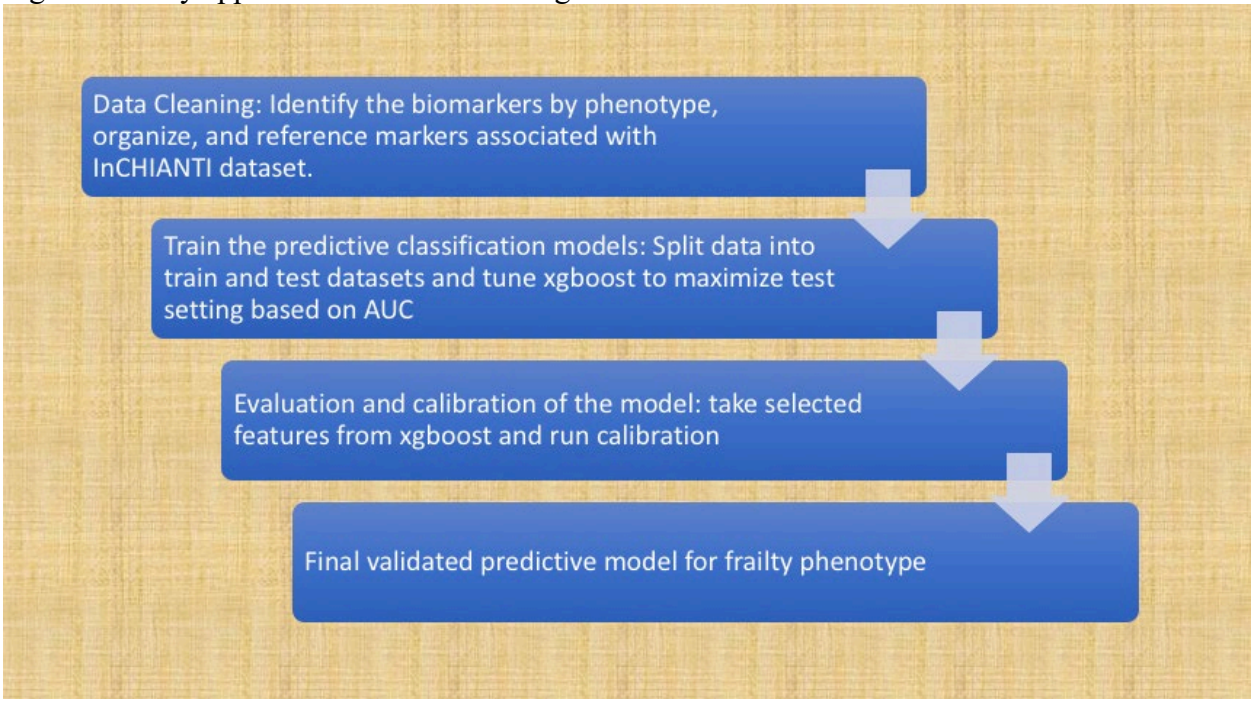
Despite evidence that frailty screening effectively identifies patients at highest risk for adverse outcomes in medical and surgical specialties, measuring frailty in clinical settings has been problematic for several reasons. After over twenty years of research, there is no universally accepted reference standard nor have predictive biological markers been established to guide clinicians in the early detection or prevention of frailty.[8,9] Multiple operational definitions have been suggested, numerous functional tests, questionnaires and indexes are available which has led to confusion among clinicians and lack of utility in screening in clinical practice. Other limitations include the length of time or special equipment required to complete the frailty screening instruments which can be a hindrance to providers who are under pressures to maintain high patient productivity.[10] A critical barrier to advancing the science has been a lack of understanding the root causes for the physiological and functional changes that define the frailty phenomenon.[11]

5. Main Hypothesis/Study Questions:
We propose to replicate a cross-setional predictive model of frailty status that was developed in the InCHIANTI study within the ARIC and Health ABC population. A predictive model that incorporates the top predictive clinical and laboratory measures would be easy to administer in a busy clinical setting. Additionally, it will provide insight into the biological mechanisms that predict a patient's potential to transition from a pre-frail to frail state and generate research into clinical and biologically based treatment strategies for the prevention of frailty outcomes.

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

1. Methods/Study Design
A cross-sectional predictive model comparing predictors with robust, pre-frail and frail participants developed from the baseline InCHIANTI data will be used to replicate a population predictive model in ARIC and Health ABC. Health ABC and ARIC models will be run separately in the datasets with a compairison of the AUC. The results will discuss the attributable AUC differences in part based on outcome derivations. Table I represents the top predictive biological features for the pre-frail and frail model 1 from the InCHIANTI study with AUC 0.9387 (95% CI 0.89-0.98). Using data from the Atherosclerosis Risk in Communities Study (ARIC), predictors from Model 1 developed in the InCHIANTI data will be used to replicate the frailty model. Figure 1 represents the workflow process that will be used to replicate the frailty prediction model in the ARIC dataset.

Figure 1. study approach and workflow diagram



Note. Profile of model development and validation workflow. Blue boxes indicate steps of the workflow specific to the ARIC dataset.

1.1. Model generation

The predictive clinical and laboratory biomarkers were identified in InCHIANTI Model 1 will be analyzed using an Extreme Gradient Boosting (xgboost) in R[12] for the validation model in ARIC. While boosting was initially developed for machine learning, 'xgboost' in R is based in boosted trees. Xgboost is an open source tool and a variant of the gradient boosting machine and uses a tree based model. Xgboost is used in this study for a supervised learning problem where the variables identified from the systematic review are used to predict pre-frail and frail individuals.

1.2. Evaluation of the model

With the use of any predictive model in machine learning there is a chance for inflated risk of capitalizing on chance features (overfitting) in the data. Overfitting of the model will be mitigated in two ways: 1) having a distinct training and validation process for the model and 2) using xgb in R which has built-in parameter settings for selection to reduce poor predictive performance. *Internal validation:* A randomly assigned training subset will be used to validate the model within the ARIC cohort *in silico* (via simulation).

1.3. Calibration of the model

Parameter estimates for each predictive factor and associated descriptive statistics will be evaluated to provide biological insight into the underpinnings of the classification algorithm. We will first evaluate the calibration by partitioning the data into 5, 10, 20, 30, 40, 50, 75, 100 and 200 groups and then run the calibration test. Next, we will repeat tests for all possible values between 5-200 groups and evaluated the distribution of the test statistic. The best prediction thresholds will be determined using AUC.

1.4. Phenotype

The frailty phenotype will be defined in three categories—non-frail (0), pre-frail(1-2), and frail (3-5).[3,13] Outcome measure of frailty will be used from V5 in the ARIC dataset.

1.5. Predictors

1.5.1. Anticholinergic Burden Calculation

The Anticholinergic Cognitive Burden (ACB) scale is the most validated scale for evaluating adverse health outcomes including cognitive and physical function[14,15]. The anticholinergic properties of each medication will be quantified using the ACB scale based on each drug's serum anticholinergic activity[16]. To determine ACB scores, each participants' medications will be assigned points (0, 1, 2, 3) according to the published 2012 update and summed for a total anticholinergic burden score. Higher scores indicate higher anticholinergic properties. An example of medications with ACB scores include: Amitriptyline = 3, Amantadine = 2, and Atenolol = 1.

1.5.2. Depression Score

The CES-D self-report scale (0-60) is used to measure depressive symptoms. Reliability, validity, and factor structure have been similar across a diverse demographic and the scale has been used extensively in epidemiologic studies for depression and physical function.[17] The CES-D score will be used in the predictive model.

1.5.3 Demographics

Age – used as a continuous variable, race/ethnicity, and education level

1.5.4. Table I. represents the validated biological and clinical markers used to predict frailty in the InCHIANTI data. Table II. Describes the clincal and biological markers available for the prediction of frailty in the ARIC dataset. A total of 23 out of the original 30 clincal and biological markers in Model I – InCHAIANTI are available in ARIC.

Table I. Biological and Clinical Predictors Model I in InCHIANTI

Inflammatory/Immunity	**Endocrine/Hormones**	**Renal/Electrolyte**
24-hour urinary cortisol (µg/24 hours)	25(OH)-D (25-hydroxyvitamin D) via RIA (nmol/L)	Creatine phosphokinase (U/L)
Erythrocyte sedimentation rate (ESR) (mm/hour)	Free testosterone (ng/dL), Vermeulen	Creatinine clearance, 24-hr urine (mL/minute)
Homocysteine via FPIA analysis (Âµmol/L)	Blood glucose (mg/dL)	Urine proteins (mg/dL)
Interleukin-1B via ELISA (pg/mL)	Free thyroxine, fT4 (ng/dL)	24-hour urinary creatinine (mg/24 hours)
Interleukin-6 via ELISA ultrasensitive (pg/mL)	Parathyroid hormone, two-site immunoradiometric assay (pg/mL)	Blood urea nitrogen (mg/dL)
Monocyte chemoattractant protein-1 via Bio-Plex (pg/mL)	**Metabolomics(plasma lipids)**	**Nutrient Biomarker**
Soluble TNF-a receptor I via quantitative sandwich EIA (pg/mL)	Fatty acid C24:0 weight (mg/L)	Vitamin B6 via high performance liquid chromatography (ng/mL)
Soluble TNF-a receptor II via quantitative sandwich EIA (pg/mL)	Lipids: HDL cholesterol (mg/dL)	Vitamin E gamma tocopherol, high performance liquid chromatography (Âµmol/L)
Hematology/Liver		Lycopene via high performance liquid chromatography (Âµmol/L)
Folate via RIA (ng/mL)		**Clinical**
Mean corpuscular volume (MCV) (fL)		Age
Retinol via high performance liquid chromatography (µmol/L)		Anticholinergic Burden
GPT (also known as ALT) (U/L)		Depression/CES-D self-report scale

Table II. Available Biological and Clinical predictors for Frailty Validation in ARIC

Inflammatory/Immunity	**Time Point**	**Renal/Electrolyte**	**Time Point**	**Clinical**
Homocysteine via FPIA analysis	V1	Creatinine clearance, 24-hr urine	V5 (n=6,493)	List of Medications for ACB score
Interleukin-1B via ELISA	V1 AS#2009.18	24-hour urinary creatinine (mg/24 hours)	V5 (n=6,493)	CES-D self-report scale
Interleukin-6 via ELISA ultrasensitive	V1 & V2 - AS#2009.18 & 1995.09	Creatine	V4, V5 AS#2009.02	Age
Monocyte chemoattractant protein-1	V1	Blood urea nitrogen	V1 (n=6,489)	
Soluble TNF-a receptor	V4	**Nutrient Biomarker**		
Hematology/Liver		Vitamin B6	V3 (n=6,191)	
Folate via RIA	V3 (n=6,191)	Vitamin E gamma tocopherol	V3 (n=6,191)	
Mean corpuscular volume (MCV)	V3 (n=6,191)			
GPT (also known as ALT)	V4 (n=6,538)			
Endocrine/Hormones				
25(OH)-D (25-hydroxyvitamin D)	V2 (n=6191)			
Testosterone	V4 (n=6,394) AS#2013.20 and AS#2013.21			
Blood glucose	V5 (n=6,495)			
Free thyroxine, fT4	V5 (n=6,441)			
Parathyroid hormone	V2 AS#2009.16			
Metabolomics(plasma lipids)				
Lipids: HDL cholesterol	V5 (n=6,538)			

1.6 Potential Limitations

1. Several of the biomarker measurements come from different time points than the outcome measure of frailty at V5. The model will be built with data as close to the outcome diagnosis (V5 Frailty) as possible; model will use data closest to visit 5 to examine AUC but also examine

model parameters and AUC adding variables from visits 1 through 4. As varibles are added, parameters (model fit and AUC) will be examined for best fit. Initial rebuild in InCHIANTI with the predictors available in ARIC at V3-V5 maintained an AUC of 93%. Although we cannot completely control for the varying temporal differences between some predictors and the frailty outcome, we anticipate findings will be informative and useful for future funding opportunities.
2. Frailty measure differs slightly between databases. Often there are vaiations in how fraity is measured between longitudinal studies which may affect the AUC and the ability of the InCHIANTI model to "fit" or accurately predicted frailty in the ARIC data. However, frailty in ARIC was based on the same frailty phenotype that is used in InCHIANTI and has been validated (key authors included on this proposal), and we do not anticipate meaningful differences in relations of biomarkers to frailty across the studies.
3. Demographics and ethic/racial differences between databases. The InCHIANTI databases is a geographically homogeneous population with a large white/caucation Euopean population. We consider it a strength rather than a limitation that ARIC and Health ABC are more diverse populations and will allow external validation to be conducted in different cohorts. Models will explore ethic/racial differences in the model fit as the model is being built. It is unknown until anaysis if the fit of the model will be dependent on these factors.

7.a. Will the data be used for non-CVD analysis in this manuscript? _x_ Yes ____ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used? _x_ Yes ____ No
(This file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ____ Yes ___X_ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"? ____ Yes ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.
ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/ARIC/search.php

__X____ Yes _____ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?
Ancilary studies used will include the investigator of this study on the publication.
MP 2465 Operationalizing frailty in the ARIC cohort (Kucharska-Newton, Palta, Windham et al)

MP 2930 Systemic inflammation in midlife as a predictor of frailty in late-life: The ARIC Study (Walker, Walston, Gottesman, Windham)
MP 2791 Association of Life's Simple 7 at Mid-life with Frailty in Older Adults. (Kucharska-Newton, Palta, Windham et al)

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __X__ Yes ____ No

11.b. If yes, is the proposal

___ A. primarily the result of an ancillary study (list number* _____)
___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* 2009.18 & 1995.09, 2013.21, 2013.20, 2009.02, 2009.16

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

References

1. Vincent GK, Velkoff VA. *THE NEXT FOUR DECADES The Older Population in the United States: 2010 to 2050 Population Estimates and Projections Current Population Reports*.; 2010. www.census.gov/population/www. Accessed November 2, 2018.
2. Hall MJ, DeFrances CJ, Williams SN, Golosinskiy A, Schwartzman A. National Hospital Discharge Survey: 2007 summary. *Natl Health Stat Report*. 2010;(29):1-20, 24. http://www.ncbi.nlm.nih.gov/pubmed/21086860. Accessed April 18, 2018.
3. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001;56(3):M146-56. http://www.ncbi.nlm.nih.gov/pubmed/11253156. Accessed August 27, 2014.
4. Makary MA, Segev DL, Pronovost PJ, et al. Frailty as a Predictor of Surgical Outcomes in Older Patients. *J Am Coll Surg*. 2010;210(6):901-908. doi:10.1016/j.jamcollsurg.2010.01.028
5. Mocchegiani E, Malavolta M, Lattanzio F, et al. Cu to Zn ratio, physical function, disability, and mortality risk in older elderly (ilSIRENTE study). *Age (Omaha)*. 2012;34(3):539-552. doi:10.1007/s11357-011-9252-2
6. Global Health and Aging. http://www.nia.nih.gov/sites/default/files/global_health_and_aging.pdf. Accessed February 8, 2015.
7. Mohanty S, Rosenthal RA, Russell MM, Neuman MD, Ko CY, Esnaola NF. Optimal

Perioperative Management of the Geriatric Patient: A Best Practices Guideline from the American College of Surgeons NSQIP and the American Geriatrics Society. *J Am Coll Surg*. 2016;222(5):930-947. doi:10.1016/j.jamcollsurg.2015.12.026
8. Panza F, Solfrizzi V, Frisardi V, et al. Different models of frailty in predementia and dementia syndromes. *J Nutr Health Aging*. 2011;15(8):711-719. http://www.ncbi.nlm.nih.gov/pubmed/21968870. Accessed February 5, 2015.
9. Bergman H, Ferrucci L, Guralnik J, et al. Frailty: an emerging research and clinical paradigm--issues and controversies. *J Gerontol A Biol Sci Med Sci*. 2007;62(7):731-737. http://www.ncbi.nlm.nih.gov/pubmed/17634320. Accessed November 19, 2017.
10. Munyon R, Cohn SL, Slawski B, Smetana GW, Pfeifer K. 2017 Update in perioperative medicine: 6 questions answered. *Cleve Clin J Med*. 2017;84(11):863-872. doi:10.3949/ccjm.84a.17068
11. Mohler MJ, Fain MJ, Wertheimer AM, Najafi B, Nikolich-Žugich J. The Frailty Syndrome: Clinical measurements and basic underpinnings in humans and animals. *Exp Gerontol*. 2014;54. doi:10.1016/j.exger.2014.01.024
12. Chen, Tianqi, He, Tong, Benesty, Michael, Khotilovich, Vadim, Tang Y. 'xgboost'- Extreme Gradient Boosting. 2017. doi:10.1145/2939672.2939785>
13. Kucharska-Newton AM, Palta P, Burgard S, et al. Operationalizing Frailty in the Atherosclerosis Risk in Communities Study Cohort. *Journals Gerontol Ser A Biol Sci Med Sci*. 2016;72(3):glw144. doi:10.1093/gerona/glw144
14. Salahudeen MS, Duffull SB, Nishtala PS. Anticholinergic burden quantified by anticholinergic risk scales and adverse outcomes in older people: a systematic review. 2015. doi:10.1186/s12877-015-0029-9
15. Campbell NL, Boustani M, Lane K, et al. Use of anticholinergics and the risk of cognitive impairment in an African-American population. *Neurology*. 2010;75:152-159. http://ovidsp.ovid.com/ovidweb.cgi?T=JS%7B&%7DPAGE=reference%7B&%7DD=emed11%7B&%7DNEWS=N%7B&%7DAN=70990267.
16. Cai X, Campbell N, Khan B, Callahan C, Boustani M. Long-term anticholinergic use and the aging brain. *Alzheimer's Dement*. 2013;9(4):377-385. doi:10.1016/j.jalz.2012.02.005
17. Radloff LS. A Self-Report Depression Scale for Research in the General Population. *Appl Psychol Meas*. 1977;1(3):385-401. doi:10.1177/014662167700100306